# FORMS REPRESENTING FORMS: THE DEFINITE CASE

JULIA BRANDES

ABSTRACT. Let $\psi$ and $F$ be positive definite forms with integral coefficients of equal degree. Using the circle method, we establish an asymptotic formula for the number of identical representations of $\psi$ by $F$, provided $\psi$ is everywhere locally representable and the number of variables of $F$ is large enough. In the quadratic case this supersedes a recent result due to Dietmann and Harvey. Another application addresses the number of primitive linear spaces contained in a hypersurface.

## 1. INTRODUCTION

Understanding the solution sets of diophantine equations in the integers is one of the pervading themes in number theory. Heuristic arguments tell us that a homogeneous polynomial $F \in \mathbb{Z}[x_1, \ldots, x_s]$ of degree $d$ whose variables are contained in a box of sidelength $P$ will take any value between $P^d$ and $-P^d$ on average roughly $P^{s-d}$ times. This heuristic has been confirmed in a classical paper by Birch [1], provided that the equation has non-singular solutions over all local fields and the number of variables satisfies $s - \dim \operatorname{Sing} F > 2^d(d-1)$. Here $\operatorname{Sing} F$ denotes the singular locus of the variety defined by $F$.

In recent work by the author [2] we derive a multidimensional analogue of Birch's theorem by investigating the number of identical representations of one homogeneous polynomial by another. Let $F \in \mathbb{Z}[x_1, \ldots, x_s]$ and $\psi \in \mathbb{Z}[t_1, \ldots, t_m]$ be forms of equal degree $d \geq 3$ with $m \geq 2$ and

$$s - \dim \operatorname{Sing} F > 3 \cdot 2^{d-1}(d-1)(r+1), \tag{1.1}$$

where $r = \binom{m+d-1}{d}$ is the number of coefficients of $\psi$. Then the number of $m$-tuples $(\mathbf{x}_1, \ldots, \mathbf{x}_m) \in \mathbb{Z}^{ms}$ of height at most $P$ satisfying

$$F(\mathbf{x}_1 t_t + \cdots + \mathbf{x}_m t_m) = \psi(t_1, \ldots, t_m) \tag{1.2}$$

identically in $t_1, \ldots, t_m$ is given by $P^{ms-rd}(c + o(1))$, where $c$ is a non-negative constant encoding the local behaviour of the problem and depending on $F$, $\psi$ and, crucially, $P$. Whilst this result is adequate in the case when $F$ and $\psi$ are both indefinite, in the definite case it fails to capture the natural size constraints on the variables $\mathbf{x}_1, \ldots, \mathbf{x}_m$ as imposed by the target polynomial $\psi$.

The objective of the present work is to address this weakness and derive an asymptotic formula for the number of representations of one positive definite form by another that reflects the correct order of magnitude as determined by the extremal dimensions of $\psi$. Suppose that $F \in \mathbb{Z}[x_1, \ldots, x_s]$ and $\psi \in \mathbb{Z}[t_1, \ldots, t_m]$ are positive definite forms of degree $d$, and that $\psi$ is given by

$$\psi(t_1, \ldots, t_m) = \sum_{\mathbf{j} \in J} n_{\mathbf{j}} t_{j_1} \cdot \ldots \cdot t_{j_d},$$

where the multi-indices $\mathbf{j} = (j_1, \ldots, j_d)$ run over the set $J = \{1, \ldots, m\}^d$ disregarding order, so that $\operatorname{Card} J = r$. For brevity, we denote the diagonal coefficients as $n_{i,\ldots,i} = n_i$. We are interested in the number

$$N(F; \psi) = \operatorname{Card}\{\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{Z}^s : F(\mathbf{x}_1 t_1 + \cdots + \mathbf{x}_m t_m) = \psi(t_1, \ldots, t_m)\}$$

of identical representations of $\psi$ by $F$. We need the magnitude $\langle\psi\rangle = \prod_{i=1}^m n_i$ of $\psi$ and its eccentricity

$$\mathcal{E}(\psi) = \max_{1 \le j \le m} \frac{\log\langle\psi\rangle}{\log n_j^m}.$$

Then our main result is as follows.

**Theorem 1.1.** *Let $F \in \mathbb{Z}[x_1, \ldots, x_s]$ and $\psi \in \mathbb{Z}[t_1, \ldots, t_m]$ be positive definite forms of equal degree $d \ge 2$, and let $m \ge 2$. Furthermore, assume that*

$$s - \dim \operatorname{Sing} F > 2^{d-1} \max\{2r(d-1), rd\,\mathcal{E}(\psi)\}.$$

*Then for some $\delta > 0$ we have*

$$N(F; \psi) = \langle\psi\rangle^{\frac{ms-rd}{md}} \chi_\infty(F; \psi) \prod_{p \text{ prime}} \chi_p(F; \psi) + O\left(\langle\psi\rangle^{\frac{ms-rd}{md}-\delta}\right),$$

*where the Euler product converges and the factors $\chi_\infty(F; \psi)$ and $\chi_p(F; \psi)$ are positive if the identity* (1.2) *has non-singular solutions over $\mathbb{R}$ and over $\mathbb{Q}_p$, respectively.*

Theorem 1.1 allows us in a very general fashion to describe the number of representations of one definite form by another. Whilst the main result of [2], which we improve slightly, points in the same direction, our previous result fails to correctly track the dimensions of the target polynomial, instead encoding them implicitly in the density of real solutions. Here we make this dependence explicit for a large class of forms, capturing the notion that the natural height conditions of the vectors $\mathbf{x}_i$ are determined by the extremal dimensions of $\psi$ as determined by its diagonal coefficients. Heuristically, after expanding (1.2) in powers of $t_1, \ldots, t_m$ and comparing coefficients one sees that the height of each $\mathbf{x}_i$ is naturally bounded by roughly $n_i^{1/d}$. Accordingly, $\psi$ will not be represented if for some $n_{\mathbf{j}}$ one has $n_{\mathbf{j}} \gg (n_{j_1} \cdot \ldots \cdot n_{j_d})^{1/d}$. We thus call $\psi$ *pseudo-diagonal* if $n_{\mathbf{j}} \le (n_{j_1} \cdot \ldots \cdot n_{j_d})^{1/d}$ for every $\mathbf{j} \in J$. It follows from the theorem that the property of being pseudo-diagonal is invariant under linear change of variables since obviously our counting function $N(F; \psi)$ is not affected by such transformations. It is also not hard to see (see Lemma 2.1) that every positive definite quadratic form is automatically pseudo-diagonal, but it is not clear to the author whether there is an easy characterisation of pseudo-diagonality within the set of positive definite forms of higher degrees. However, we will see that the real solution density is zero if $\psi$ is not close to being pseudo-diagonal, but when it is, Theorem 1.1 shows that the growth is determined by the diagonal contribution while the factor $\chi_\infty(F; \psi)$ provides a correction factor which characterises the deformity of $\psi$ arising from the off-diagonal contributions.

One aspect of Theorem 1.1 that strikes the eye is the dependance of the result on the eccentricity $\mathcal{E}(\psi)$. One sees that $\mathcal{E}(\psi) = 1$ when $n_1 = \cdots = n_m$, while it will be large if the diagonal entries vary significantly, so $\mathcal{E}(\psi)$ measures how far the body described by $\psi(\mathbf{t}) \le 1$ is stretched or contracted away from the hypercube. This reflects the fact that if one of the $n_i$ is very small, the respective variable $\mathbf{x}_i$ is essentially fixed, which would profoundly change the character of the problem as

the polynomials would cease to be homogeneous and of the same degree.

Prior to this, similar results concerning the representation of definite forms had only been available in the quadratic case, where a wider range of methods is available. In fact, an asymptotic formula for the number $N(A; B)$ of identical representations of a quadratic form $B \in \mathrm{Mat}(\mathbb{Z}, m)$ by a positive definite quadratic form $A \in \mathrm{Mat}(\mathbb{Z}, s)$ has already been known since Raghavan's work of the late 1950s [6]. We may assume without loss of generality that the matrices $A$ and $B$ are non-singular. It follows from the pseudo-diagonality of $B$ that $|b_{i,j}| \leq b_j$ whenever $b_i < b_j$, so $B$ is essentially Minkowski reduced. Raghavan's result states that for all matrices $A$ and $B$ as above with $s \geq 2m + 3$ and $b_i \gg (\det B)^{1/m}$ for all $1 \leq i \leq m$ there exists some positive constant $\delta$ such that

$$N(A; B) = c_{s,m}(\det A)^{-m/2}(\det B)^{\frac{s-m-1}{2}} \prod_{p \text{ prime}} \chi_p(A; B) + O\left(\det B^{\frac{s-m-1}{2}-\delta}\right),$$
(1.3)

where

$$c_{s,m} = (\sqrt{\pi})^{ms-\frac{m(m+1)}{2}} \prod_{j=m+1}^{s} \Gamma\left(\frac{j-m}{2}\right)^{-1}$$
(1.4)

and

$$\chi_p(A; B) = \lim_{l \to \infty} (p^l)^{\frac{m(m+1)}{2} - ms} \mathrm{Card}\left\{X \pmod{p^l} : X^t A X \equiv B \pmod{p^l}\right\}.$$

More recently, Dietmann and Harvey [5] showed via the circle method that (1.3) is true whenever

$$s > 2\left(\frac{m(m+1)}{2} + 1\right)\left(\frac{m(m+1)}{2} + \max_{1 \leq j \leq m} \sum_{i=1}^{m} \frac{\log(b_i/b_j)}{\log b_j}\right).$$

Whilst this condition is vastly more restrictive on the number of variables than that of Raghavan, it dispenses with the latter's strict dependence on the relative sizes of the diagonal entries of $B$. However, Theorem 1.1 allows us to strengthen their result.

**Corollary 1.** *Suppose $A$ and $B$ are as above, with*

$$s > (m + 1)\max_{1 \leq j \leq m} \sum_{i=1}^{m} \frac{\log b_i}{\log b_j}.$$

*Then* (1.3) *is satisfied.*

One observes that, while our bound is still far from Raghavan's, it supersedes that of Dietmann and Harvey by a factor $m^2$ whilst retaining the flexibility with respect to the diagonal entries of $B$. This improvement has been made possible by the advances made in [2] regarding the multidimensional version of Birch's theorem.

In the case $\psi = 0$ Equation (1.2) describes an $m$-dimensional linear space on the hypersurface defined by $F = 0$. Indeed, since linear spaces over the integers can be interpreted as lower-dimensional sublattices of $\mathbb{Z}^s$, our work in [2] shows that the number of $m$-dimensional lattices $X \in \mathbb{Z}^{s \times m}$ with generators of height at most $P$ on which $F$ vanishes identically grows like $P^{ms-rd}$, provided (1.1) is satisfied and

the problem is everywhere locally soluble. We may now generalise this question and ask for the number of $m$-dimensional lattices $X$ contained in a non-trivial sublattice of $\mathbb{Z}^{s \times m}$. Thus for a given matrix $C \in \mathrm{Mat}(\mathbb{Z}, m)$ we are interested in the number of lattices $X \in \mathbb{Z}^{s \times m} C$ with generators of length at most $P$ on which $F$ vanishes identically. Such questions arise naturally when one tries to implement an inclusion-exclusion argument on the set of lattices in order to restrict the count to primitive linear spaces.

**Theorem 1.2.** *Let $d \geq 2$ and $m \geq 2$ be positive integers, and let $P$ be large. Further, let $C \in \mathbb{Z}^{m \times m}$ be a non-singular matrix whose Smith normal form is given by $\mathrm{diag}(\gamma_1, \ldots, \gamma_m)$. Then, provided that*

$$s - \dim \mathrm{Sing}\, F > 2^{d-1} \max \left\{ 2r(d-1), rd \frac{\sum_{i=1}^{m} \log(P/\gamma_i)}{m \log(P/\gamma_{\max})} \right\},$$

*the number $N_C(P)$ of points $(\mathbf{x}_1, \ldots, \mathbf{x}_m) \in \mathbb{Z}^{s \times m} C$ of height at most $P$ solving*

$$F(\mathbf{x}_1 t_t + \cdots + \mathbf{x}_m t_m) = 0 \tag{1.5}$$

*identically in $t_1, \ldots, t_m$ is given by*

$$N_C(P) = \left( \frac{P^m}{\det C} \right)^s P^{-rd} \chi_\infty \prod_{p \text{ prime}} \chi_p + o\left( \left( \frac{P^m}{\det C} \right)^s P^{-rd} \right),$$

*and $\chi_p$ and $\chi_\infty$ characterise the local solution densities of the variety defined by $F = 0$ and are independent of $C$.*

Observe that in the case $m = 1$ Theorem 1.2 reduces to counting solutions $\mathbf{x} \in (d\mathbb{Z})^s$ of height $P$ solving $F(\mathbf{x}) = 0$, which by homogeneity is equivalent to counting $\mathbf{x} \in \mathbb{Z}^s$ of height $P/d$. In the higher-dimensional setting, however, the situation is more complicated, but Theorem 1.2 shows that nonetheless a similar argument can be made even in this case. This situation is generalised for higher values of $m$ here. Of course, a weakness of Theorem 1.2 is that it does not account for unimodular coordinate transforms. This is an issue we hope to address in future work.

Meanwhile, in the case $C = \mathrm{Id}$ the result in Theorem 1.2 as well as Theorem 1.1 improves on our former result in [2], replacing the condition $s - \dim \mathrm{Sing}\, F > 3 \cdot 2^{d-1}(d-1)(r+1)$ by the milder $s - \dim \mathrm{Sing}\, F > 2^d r(d-1)$. This improvement carries over to the case of systems of $R$ forms, where the respective bound on the number of variables is given by $s - \dim \mathrm{Sing}\, \mathbf{F} > 2^{d-1}(d-1)Rr(R+1)$.

## 2. Setup

**Notation.** Although we believe our notation to be largely self-explanatory, we would like to point out a few items that will be used recurringly.

We will use the Vinogradov and Landau symbols throughout. Whenever the letter $\varepsilon$ occurs, the respective statement is true for all $\varepsilon > 0$. We will therefore not trace the particular 'value' of each $\varepsilon$, which can consequently change from statement to statement. Furthermore, whenever we write $\sum_{n=a}^{b} f(n)$ with possibly non-integral

boundaries $a$ or $b$, the sum is to be understood to mean $\sum_{a \leq n \leq b} f(n)$. Occasionally we will write $\sum_{x \ll P} f(x)$, which should be interpreted as $\sum_{-c_1 P \leq x \leq c_2 P} f(x)$ with suitable absolute constants $c_1, c_2 > 0$.

We will abuse vector notation extensively, so any statement involving vectors should be read entry-wise. In this vein we will write $|\mathbf{x}| \leq P$ to mean $|x_i| \leq P$ for all entries $i$. Similarly, $(a, \mathbf{x})$ denotes the greatest common divisor $(a, x_1, \ldots, x_n)$ of $a$ and all entries of $\mathbf{x}$. We are confident that no misunderstandings will arise if all similar statements are read in a like manner.

Finally, we will occasionally write $P_{\min} = \min_i P_i$, and similarly $\gamma_{\max} = \max_i \gamma_i$.

Before embarking on the proof of our results, it is useful to show how they are connected with one another. To enhance clarity, we reformulate Theorem 1.1 as follows.

**Theorem 2.1.** *Let $F \in \mathbb{Z}[x_1, \ldots, x_s]$ and $\psi \in \mathbb{Z}[t_1, \ldots, t_m]$ be forms of degree $d \geq 2$, where $m \geq 2$, and let $P_1 \leq \cdots \leq P_m$ be large. Furthermore, suppose that*

$$s - \dim \operatorname{Sing} F > 2^{d-1} \max \left\{ 2r(d-1), rd \left( \frac{\sum_{i=1}^m \log P_i}{m \log P_1} \right) \right\}.$$

*Then the number $N_\psi(P_1, \ldots, P_m)$ of points $\mathbf{x}_i \in [-P_i, P_i]^s \cap \mathbb{Z}^s$ $(1 \leq i \leq m)$ solving (1.2) identically in $t_1, \ldots, t_m$ is given by*

$$N_\psi(P_1, \ldots, P_m) = \left( \prod_{i=1}^m P_i \right)^{s - rd/m} \chi_\infty(F; \psi) \prod_{p \text{ prime}} \chi_p(F; \psi) + o\left( \left( \prod_{i=1}^m P_i \right)^{s - rd/m} \right),$$

*where the factors are given by*

$$\chi_\infty(F; \psi) = \operatorname{vol}_\infty \left\{ |\boldsymbol{\xi}_i|_\infty \leq 1 \ (1 \leq i \leq m) : F(\boldsymbol{\xi}_1 t_1 + \cdots + \boldsymbol{\xi}_m t_m) = \psi(\tfrac{t_1}{P_1}, \ldots, \tfrac{t_m}{P_m}) \right\}$$

*and*

$$\chi_p(F; \psi) = \operatorname{vol}_p \{ |\boldsymbol{\xi}_i|_p \leq 1 \ (1 \leq i \leq m) : F(\boldsymbol{\xi}_1 t_1 + \cdots + \boldsymbol{\xi}_m t_m) = \psi(t_1, \ldots, t_m) \}$$

*and denote the $(ms - r)$-dimensional volume of the normalised set of solutions in the real and $p$-adic unit cubes, respectively.*

Theorem 1.1 follows from here by setting $P_i = n_i^{1/d}$. With this choice, the dependence on $P_1, \ldots, P_m$ of the real solution density amounts to a re-normalisation of the target form to unit length, where $\psi$ is replaced by

$$\widetilde{\psi}(t_1, \ldots, t_m) = \sum_{\mathbf{j} \in J} \tilde{n}_{\mathbf{j}} t_{j_1} \cdot \ldots \cdot t_{j_d} \tag{2.1}$$

with coefficients $\tilde{n}_{\mathbf{j}} = n_{\mathbf{j}} (n_{j_1} \cdot \ldots \cdot n_{j_d})^{-1/d}$ which satisfy $\tilde{n}_{\mathbf{j}} \leq 1$ for all $\mathbf{j} \in J$ whenever $\psi$ is pseudo-diagonal. Thus, while the size of the main term is determined by the absolute dimensions of $\psi$ as defined by $\langle \psi \rangle$, the real solution density provides a correction factor by tracking the intrinsic deformity of the body defined by $\psi$ that is preserved after its renormalisation to unit length.

Similarly, Corollary 1 follows from specialising $d = 2$, so that the form $B$ is described by an $(m \times m)$-matrix. In the quadratic case the question of whether or not a positive definite form is pseudo-diagonal is easy to answer.

**Lemma 2.1.** *Every positive definite quadratic form is pseudo-diagonal.*

*Proof.* Obviously every positive definite matrix has non-negative diagonal entries, for if $\mathbf{e}_i$ denotes the $i$-th unit vector, one has $b_i = \mathbf{e}_i^t B \mathbf{e}_i \geq 0$. Suppose now that $B$ is a symmetric matrix that is not pseudo-diagonal, so for some $i, j$ one has $|b_{ij}| > \sqrt{b_i b_j}$. Let us first consider the case $b_{ij} > \sqrt{b_i b_j}$, then for every choice of $\lambda, \mu > 0$ one has

$$(\lambda \mathbf{e}_i - \mu \mathbf{e}_j)^t B (\lambda \mathbf{e}_i - \mu \mathbf{e}_j) = \lambda^2 b_i - 2\lambda\mu b_{ij} + \mu^2 b_j$$
$$< \lambda^2 b_i - 2\lambda\mu \sqrt{b_i b_j} + \mu^2 b_j = (\lambda \sqrt{b_i} - \mu \sqrt{b_j})^2.$$

Choosing $\lambda = \sqrt{b_j}$ and $\mu = \sqrt{b_i}$ delivers a contradiction to the assumption that $B$ is positive definite. On the other hand, if $b_{ij} < -\sqrt{b_i b_j}$ we consider the expression

$$(\lambda \mathbf{e}_i + \mu \mathbf{e}_j)^t B (\lambda \mathbf{e}_i + \mu \mathbf{e}_j) = \lambda^2 b_i + 2\lambda\mu b_{ij} + \mu^2 b_j$$
$$< \lambda^2 b_i - 2\lambda\mu \sqrt{b_i b_j} + \mu^2 b_j = (\lambda \sqrt{b_i} - \mu \sqrt{b_j})^2,$$

where again we assumed $\lambda, \mu > 0$, and the remainder of the argument follows as above. This shows the statement. $\square$

It follows that for any positive definite matrix one has $\det B \asymp b_1 \cdot \ldots \cdot b_m$, and in fact, whilst from Theorem 2.1 we obtain

$$\chi_\infty(A; B) = \int_{\mathbb{R}^{m(m+1)/2}} \int_{[-1,1]^{ms}} e\left( \beta_{ij} \left( \boldsymbol{\xi}_i^T A \boldsymbol{\xi}_j - \frac{b_{ij}}{\sqrt{b_i b_j}} \right) \right) d\boldsymbol{\xi}_1 \cdots d\boldsymbol{\xi}_m d\boldsymbol{\beta},$$

(see equation (4.8) below), it follows from [5, §6] that indeed

$$(b_1 \cdot \ldots \cdot b_m)^{(s-m-1)/2} \chi_\infty(A; B) = (\det A)^{-m/2} (\det B)^{(s-m-1)/2} c_{s,m},$$

where $c_{s,m}$ is as given in (1.4).

Also the third result given in the introduction is essentially a special case of Theorem 2.1 corresponding to the zero polynomial $\psi = 0$. First, observe that Theorem 1.2 depends only on the Smith normal form of $C$. This can be seen as follows. Suppose $U, V \in \mathrm{SL}_m(\mathbb{Z})$ are such that $UCV$ is diagonal. By interpreting the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$ as a matrix $X$, we can view the function $N_C(P)$ as counting matrices $X \in \mathbb{Z}^{m \times s} C$ of height at most $P$ for which $F(X\mathbf{t}) = 0$ is true identically in $\mathbf{t}$. This implies, however, that we may replace $\mathbf{t}$ by $V\mathbf{t}$ in the statement and equivalently demand that $F(XV\mathbf{t}) = 0$ identically in $\mathbf{t}$. On the other hand, the condition $X \in \mathbb{Z}^{m \times s} C$ can be written as $X = YC$ with $Y \in \mathbb{Z}^{m \times s}$, so $N_C(P)$ counts the matrices $Y$ for which $YC$ is of height at most $P$ and $F(YC\mathbf{t}) = 0$ identically in $\mathbf{t}$. Since $U$ is unimodular, we may equivalently set $Y = ZU$ and count matrices $Z \in \mathbb{Z}^{m \times s}$ such that $ZUC$ is of height at most $P$ and $F(ZUC\mathbf{t}) = 0$ identically in $\mathbf{t} \in \mathbb{Z}^m$. It follows that we may assume without loss of generality that $C$ is of the form $C = \mathrm{diag}(\gamma_1, \ldots, \gamma_m)$, and the condition

$$(\mathbf{x}_1, \ldots, \mathbf{x}_m) \in \mathbb{Z}^{s \times m} C \cap [-P, P]^{s \times m}$$

translates into the simpler

$$\mathbf{x}_i \in \gamma_i \mathbb{Z}^s \cap [-P, P]^s \qquad (1 \leq i \leq m).$$

It thus remains to relate the counting function

$$N_C(P) = \mathrm{Card}\{\mathbf{x}_i \in \gamma_i \mathbb{Z}^s \cap [-P, P]^s \ (1 \leq i \leq m) : F(\mathbf{x}_1 t_1 + \cdots + \mathbf{x}_m t_m) = 0\}$$

to the more familiar

$$N(P/\gamma_1, \ldots, P/\gamma_m) = \text{Card}\{|\mathbf{x}_i| \leq P/\gamma_i \ (1 \leq i \leq m) : F(\mathbf{x}_1 t_1 + \cdots + \mathbf{x}_m t_m) = 0\}$$

which is addressed in Theorem 1.1.

After expanding, the form $F$ may be written as

$$F(t_1\mathbf{x}_1 + \ldots + t_m\mathbf{x}_m) = \sum_{\mathbf{j} \in J} A(\mathbf{j}) t_{j_1} t_{j_2} \cdot \ldots \cdot t_{j_d} \Phi(\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \ldots, \mathbf{x}_{j_d}),$$

where we use the notation introduced in [2] in writing $\Phi$ for the symmetric $d$-linear form associated to $F$ and $A(\mathbf{j})$ for the combinatorial factors that take into account the multiplicity of each term. Furthermore, write $\overline{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ and $\boldsymbol{\alpha} = (\alpha_{\mathbf{j}})_{\mathbf{j} \in J}$, and let

$$\Phi_{\mathbf{j}}(\overline{\mathbf{x}}) = A(\mathbf{j}) \Phi(\mathbf{x}_{j_1}, \ldots \mathbf{x}_{j_d}) \qquad (\mathbf{j} \in J)$$

and

$$\mathfrak{F}(\mathbf{x}_1, \ldots, \mathbf{x}_m; \boldsymbol{\alpha}) = \sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} \Phi_{\mathbf{j}}(\overline{\mathbf{x}}). \tag{2.2}$$

In the context of Theorem 2.1 the variables $\mathbf{x}_i$ lie in intervals $[-P_i, P_i]$, the Cartesian product of which which we denote by $\mathcal{P}$. Notice that this can be transformed into the language of Theorem 1.2 by by setting $P_m = P$ and $\gamma_i = P_m/P_i$ for all $i$. In this notation we have

$$\text{Card } \mathcal{P} = \prod_{i=1}^{m} P_i = \prod_{i=1}^{m} \frac{P}{\gamma_i} = \frac{P^m}{\det C}.$$

Thus classical orthogonality relations imply that the number of solutions to (1.5) with $\mathbf{x}_i \leq P_i$ is given by

$$N_{\psi}(\mathcal{P}) = \int_{[0,1)^r} T(\boldsymbol{\alpha}; \mathcal{P}) e(-\mathbf{n} \cdot \boldsymbol{\alpha}) \mathrm{d}\boldsymbol{\alpha},$$

where the exponential sum is defined as

$$T(\boldsymbol{\alpha}; \mathcal{P}) = \sum_{\overline{\mathbf{x}} \in \mathcal{P}^s} e(\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})).$$

In the case of Theorem 1.2, on the other hand, the counting function takes the shape

$$N_C(P) = \int_{[0,1)^r} \sum_{\overline{\mathbf{x}} C \leq P} e\left(\mathfrak{F}\left(\overline{\mathbf{x}} C; \boldsymbol{\alpha}\right)\right) \mathrm{d}\boldsymbol{\alpha}.$$

Since we may assume $C$ to be diagonal, the variables $\mathbf{x}_i$ are multiples of the diagonal entries $\gamma_i$ of $C$, and by homogeneity we may write

$$\mathfrak{F}(\gamma_1\mathbf{x}_1, \ldots, \gamma_m\mathbf{x}_m; \boldsymbol{\alpha}) = \sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} A(\mathbf{j}) \Phi(\gamma_{j_1}\mathbf{x}_{j_1}, \ldots, \gamma_{j_d}\mathbf{x}_{j_d})$$

$$= \sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} \hat{\gamma}_{\mathbf{j}} A(\mathbf{j}) \Phi(\mathbf{x}_{j_1}, \ldots, \mathbf{x}_{j_d}),$$

where we introduced the notation $\hat{\boldsymbol{\gamma}}_{\mathbf{j}}$ for the product $\gamma_{j_1} \cdot \ldots \cdot \gamma_{j_d}$. Absorbing the factors $\hat{\boldsymbol{\gamma}}_{\mathbf{j}}$ into the coefficients $\alpha_{\mathbf{j}}$, we see that the number of solutions is given by

$$N_C(P) = \int_{[0,1)^r} \sum_{\overline{\mathbf{x}} \in \mathcal{P}^s} e\left(\sum_{\mathbf{j} \in J} \alpha_{\mathbf{j}} \hat{\boldsymbol{\gamma}}_{\mathbf{j}} \Phi_{\mathbf{j}}(\overline{\mathbf{x}})\right) d\boldsymbol{\alpha} = \left(\prod_{\mathbf{j} \in J} \hat{\boldsymbol{\gamma}}_{\mathbf{j}}\right)^{-1} N(\mathcal{P}).$$

The product in the last expression is symmetric in the $\gamma_i$ and has altogether $rd$ factors, so its value is

$$\prod_{\mathbf{j} \in J} \hat{\boldsymbol{\gamma}}_{\mathbf{j}} = (\gamma_1 \cdot \ldots \cdot \gamma_m)^{rd/m} = (\det C)^{rd/m}. \tag{2.3}$$

It is therefore the counting function $N_\psi(\mathcal{P})$ considered in Theorem 2.1 towards which we will direct our attention.

## 3. The tripartite Weyl inequality

The first step is to establish an inequality of Weyl type. Although the greater picture of this is by now fairly standard, the different ranges of the $\mathbf{x}_i$ create some technical complications which need to be attended to with due care.

We define the discrete difference operator by its action on the form $\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})$ as

$$\Delta_{i,\mathbf{h}}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha}) = \mathfrak{F}(\mathbf{x}_1, \ldots, \mathbf{x}_i + \mathbf{h}, \ldots, \mathbf{x}_m; \boldsymbol{\alpha}) - \mathfrak{F}(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_m; \boldsymbol{\alpha}),$$

and write for brevity

$$\Delta_{\mathbf{h_j}}^{(k)} = \Delta_{k, \mathbf{h}_{j_k}} \cdots \Delta_{1, \mathbf{h}_{j_1}}.$$

This allows us to formulate our first Weyl differencing lemma.

**Lemma 3.1.** *Let $1 \leq k \leq d-1$ and fix $\mathbf{j} \in J$. We have*

$$|T(\boldsymbol{\alpha}; \mathcal{P})|^{2^k} \ll \left(\frac{P^m}{\det C}\right)^{(2^k-1)s} \left(\prod_{l=1}^k P_{j_l}^{-s}\right) \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \leq l \leq k}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(k)} \mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right),$$

*and the sum over $\overline{\mathbf{x}}$ is over a suitable box contained in $\mathcal{P}^s$.*

*Proof.* The proof is, as usual, by induction. The case $k = 1$ follows by Cauchy–Schwarz via

$$|T(\boldsymbol{\alpha}; \mathcal{P})|^2 \ll \left(\sum_{\substack{|\mathbf{x}_i| \leq P_i \\ i \neq j_1}} 1\right) \sum_{\substack{|\mathbf{x}_i| \leq P_i \\ i \neq j_1}} \left|\sum_{|\mathbf{x}_{j_1}| \leq P_{j_1}} e(\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha}))\right|^2$$

$$\ll \left(\frac{P^m}{\det C}\right)^s P_{j_1}^{-s} \sum_{|\mathbf{h}_1| \leq P_{j_1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{j_1, \mathbf{h}_1} \mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right).$$

Note that the final summation of $\mathbf{x}_{j_1}$ is over the set

$$\{\mathbf{x}_{j_1} \in \mathbb{Z}^s : |\mathbf{x}_{j_1}| \leq P_{j_1}, |\mathbf{x}_{j_1} + \mathbf{h}_1| \leq P_{j_1}\},$$

which is again a box contained in $[-P_{j_1}, P_{j_1}]^s$. The induction step is similar. By another application of Cauchy–Schwarz one has

$$\left| \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le k}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(k)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right) \right|^2$$

$$\ll \left( \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le k}} \sum_{\substack{\mathbf{x}_i \ll P_i \\ i \ne j_{k+1}}} 1 \right) \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le k}} \sum_{\substack{\mathbf{x}_i \ll P_i \\ i \ne j_{k+1}}} \left| \sum_{\mathbf{x}_{j_{k+1}} \ll P_{j_{k+1}}} e\left(\Delta_{\mathbf{h_j}}^{(k)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right) \right|^2$$

$$\ll \left( \frac{P^m}{\det C} \right)^s P_{j_{k+1}}^{-s} \left( \prod_{l=1}^{k} P_{j_l}^s \right) \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le k+1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(k+1)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right),$$

and hence

$$|T(\boldsymbol{\alpha}; \mathcal{P})|^{2^{k+1}} \ll \left( \frac{P^m}{\det C} \right)^{(2^{k+1}-2)s} \left( \prod_{l=1}^{k} P_{j_l}^{-2s} \right) \left| \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le k}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(k)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right) \right|^2$$

$$\ll \left( \frac{P^m}{\det C} \right)^{(2^{k+1}-1)s} \left( \prod_{l=1}^{k+1} P_{j_l}^{-s} \right) \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le k+1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(k+1)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right).$$

This completes the proof. $\qquad\square$

As is usual in Weyl differencing arguments, we notice that the differencing procedure gradually reduces the degree while preserving the structure of the system. In our case, this means that after $d-1$ applications the resulting expression is linear in the variables $\mathbf{x}_1, \ldots, \mathbf{x}_m$, and since all forms in the system of equations are instances of the same multilinear form $\Phi$, all of these linear expressions will be of the same shape. We abbreviate $\mathcal{H}$ for the $(d-1)$-tuple $(\mathbf{h}_1, \ldots, \mathbf{h}_{d-1})$, and define the $(d-1)$-linear forms $B_n$, $1 \le n \le s$, via the relation

$$\Phi(\mathbf{x}, \mathcal{H}) = \sum_{n=1}^{s} B_n(\mathcal{H})x_n,$$

so that one has

$$\Delta_{\mathbf{h_j}}^{(d-1)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha}) = \sum_{j_d=1}^{m} \sum_{n=1}^{s} \alpha_{\mathbf{j}} M(\mathbf{j}) B_n(\mathcal{H}) x_{j_d,n} + R(\mathcal{H}) \tag{3.1}$$

for some combinatorial factors $M(\mathbf{j})$ and some function $R(\mathcal{H})$ collecting the terms independent of $\overline{\mathbf{x}}$. Lemma 3.1 yields now

$$|T(\boldsymbol{\alpha}; \mathcal{P})|^{2^{d-1}} \ll \left( \frac{P^m}{\det C} \right)^{(2^{d-1}-1)s} \left( \prod_{l=1}^{d-1} P_{j_l}^{-s} \right) \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le d-1}} \sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(d-1)}\mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right),$$

and by (3.1) we have

$$\sum_{\overline{\mathbf{x}}} e\left(\Delta_{\mathbf{h_j}}^{(d-1)} \mathfrak{F}(\overline{\mathbf{x}}; \boldsymbol{\alpha})\right) \ll \left(\frac{P^m}{\det C}\right)^s P_{j_d}^{-s} \left| \sum_{|\mathbf{x}_{j_d}| \ll P_{j_d}} e\left(M(\mathbf{j})\alpha_{\mathbf{j}}\Phi(\mathbf{x}_{j_d}, \mathcal{H})\right) \right|$$

$$\ll \left(\frac{P^m}{\det C}\right)^s P_{j_d}^{-s} \prod_{n=1}^{s} \min\left(P_{j_d}, \|M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H})\|^{-1}\right).$$

It follows that if we define

$$\Upsilon(\mathbf{j}) = \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 1 \le l \le d-1}} \prod_{n=1}^{s} \min\left(P_{j_d}, \|M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H})\|^{-1}\right),$$

then we can bound the exponential sum as

$$|T(\boldsymbol{\alpha}; \mathcal{P})|^{2^{d-1}} \ll \left(\frac{P^m}{\det C}\right)^{2^{d-1}s} \left(\prod_{k=1}^{d} P_{j_k}^{-s}\right) \Upsilon(\mathbf{j}). \tag{3.2}$$

We want to conclude from (3.2) that either $T(\boldsymbol{\alpha}; \mathcal{P})$ can be bounded non-trivially, or we have a good rational approximation to the vector $\boldsymbol{\alpha}$, or else the form $F$ has been highly singular from the beginning. Let $N_{\mathbf{j}}(X_1, \ldots, X_{d-1}; Y)$ denote the cardinality of the set

$$\left\{|\mathbf{h}_k| \le X_k \quad (1 \le k \le d-1) : \|M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H})\|^{-1} > Y \quad (1 \le n \le s)\right\}. \tag{3.3}$$

We show that when $T(\boldsymbol{\alpha}; \mathcal{P})$ is large, then $N_{\mathbf{j}}(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d})$ is also large.

**Lemma 3.2.** *Suppose*

$$|T(\boldsymbol{\alpha}; \mathcal{P})| \gg \left(\frac{P^m}{\det C}\right)^s P^{-K}$$

*for some parameter $K > 0$. Then for every $\mathbf{j} \in J$ one has*

$$N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right) \gg P^{-2^{d-1}K-\varepsilon} \prod_{k=1}^{d-1} P_{j_k}^{s}.$$

*Proof.* Combining the hypothesis of the lemma with (3.2), we have

$$\left(\left(\frac{P^m}{\det C}\right)^s P^{-K}\right)^{2^{d-1}} \ll \left(\frac{P^m}{\det C}\right)^{2^{d-1}s} \left(\prod_{k=1}^{d} P_{j_k}^{-s}\right) \Upsilon(\mathbf{j}).$$

Rearranging the terms, one obtains

$$\Upsilon(\mathbf{j}) \gg P^{-2^{d-1}K} \prod_{k=1}^{d} P_{j_k}^{s}. \tag{3.4}$$

Note that, unlike in the situation considered in [2], we obtain distinct estimates for different indices $\mathbf{j} \in J$.

Now for a fixed tuple $(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1})$ write

$$R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}) = \mathrm{Card}\{\mathbf{h}_1 \ll P_{j_1} : \|M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H})\| < P_{j_d}^{-1} \quad (1 \le n \le s)\},$$

so that

$$\sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 2 \le l \le d-1}} R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}) = N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right).$$

Then a familiar pigeonhole argument as in the proof of [4, Lemma 13.2] implies that for any $s$-tuple of integers $(r_1, \ldots, r_s)$ between $0$ and $P_{j_d} - 1$ and any fixed $\mathbf{h}_2, \ldots, \mathbf{h}_{d-1}$ the number of $\mathbf{h}_1 \ll P_{j_1}$ satisfying

$$\frac{r_n}{P_{j_d}} < \{M(\mathbf{j})\alpha_{\mathbf{j}} B_{\mathbf{j},n}(\mathcal{H})\} < \frac{r_n + 1}{P_{j_d}} \quad (1 \le n \le s)$$

is at most $R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1})$, and thus

$$\Upsilon(\mathbf{j}) \ll \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 2 \le l \le d-1}} \prod_{n=1}^{s} \sum_{r_n=1}^{P_{j_d}-1} \min\left\{P_{j_d}, \frac{P_{j_d}}{r_n}, \frac{P_{j_d}}{r_n + 1}\right\}$$

$$\ll (P_{j_d} \log P_{j_d})^s \sum_{\substack{\mathbf{h}_l \ll P_{j_l} \\ 2 \le l \le d-1}} R(\mathbf{h}_2, \ldots, \mathbf{h}_{d-1})$$

$$\ll (P_{j_d} \log P_{j_d})^s N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right).$$

Inserting this into (3.4) gives the desired result. $\qquad \square$

We will need the following standard lemma.

**Lemma 3.3.** *Let $L_1, \ldots, L_n$ be linear forms given by*

$$L_i = \lambda_{i,1} x_1 + \ldots + \lambda_{i,n} x_n \quad (1 \le i \le n)$$

*with the additional symmetry that $\lambda_{i,j} = \lambda_{j,i}$. For a parameter $A > 1$ let $U(Z)$ denote the number of integer solutions $x_1, \ldots, x_n$ to the system*

$$|x_i| < AZ \quad and \quad \|L_i(\mathbf{x})\| < Z/A \quad (1 \le i \le n).$$

*Then for $0 < Z' \le Z \le 1$ we have*

$$\frac{U(Z)}{U(Z')} \ll \left(\frac{Z}{Z'}\right)^n.$$

*Proof.* This is [4, Lemma 12.6]. $\qquad \square$

The strategy is now to apply Lemma 3.3 to each of the variables $\mathbf{h}_k$ in such a way that $AZ = P_{j_k}$ and $AZ' = P^\theta$ for some small exponent $\theta$, so that in the further course of the argument we can assume the variables to lie in small boxes which are then independent of $C$. However, this is legitimate only in the case when $Z' \le Z$, so we need $(AZ')/(AZ) = P^\theta/P_{j_k} \le 1$ for all $k$. This condition amounts to $P^\theta \le P/\gamma_{\max}$ or, taking logarithms,

$$\theta \le 1 - \frac{\log \gamma_{\max}}{\log P}.$$

For simplicity we write

$$\eta = \log \gamma_{\max} / \log P. \tag{3.5}$$

Notice that in the case when $C$ is the identity matrix we have $\eta = 0$ and therefore $\theta \le 1$ as usual.

**Lemma 3.4.** *Suppose that $0 < \theta \leq 1 - \eta$, where $\eta$ is as in (3.5). Then under the hypothesis of Lemma 3.2 one has*

$$N_{\mathbf{j}}\left(P^\theta, \ldots, P^\theta; \frac{P^{d-(d-1)\theta}}{\hat{\gamma}_{\mathbf{j}}}\right) \gg P^{(d-1)s\theta - 2^{d-1}K - \varepsilon}.$$

*Proof.* This follows from Lemma 3.3 by what is essentially a standard argument. We may assume without loss of generality that $\gamma_i \geq \gamma_j$ for $i < j$ and that the components of every multi-index $\mathbf{j} \in J$ are arranged in ascending order, so that $\gamma_{j_1} \geq \gamma_{j_2} \geq \cdots \geq \gamma_{j_d}$ for all $\mathbf{j} \in J$. For fixed $\mathbf{h}_1, \ldots \mathbf{h}_{k-1}, \mathbf{h}_{k+1}, \ldots, \mathbf{h}_{d-1}$ and fixed $A > 1$ let $U_k(Z)$ denote the number of $\mathbf{h}_k < AZ$ such that $\|M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H})\| < Z/A$. We will take

$$A_1 = \frac{P}{\sqrt{\gamma_{j_1}\gamma_{j_d}}}, \qquad Z_1 = \sqrt{\frac{\gamma_{j_d}}{\gamma_{j_1}}}, \qquad Z_1' = \frac{\sqrt{\gamma_{j_d}\gamma_{j_1}}}{P^{1-\theta}},$$

and then recursively

$$A_k = A_{k-1}\sqrt{\frac{P^{1-\theta}}{\gamma_{j_k}}}, \qquad Z_k = \frac{\gamma_{j_{k-1}}Z_{k-1}}{\sqrt{P^{1-\theta}\gamma_{j_k}}}, \qquad Z_k' = \frac{\gamma_{j_k}Z_k}{P^{1-\theta}},$$

so that the relations

$$A_kZ_k = P/\gamma_{j_k}, \qquad\qquad Z_k/A_k = P^{-k+(k-1)\theta}\gamma_{j_1}\cdot\ldots\cdot\gamma_{j_{k-1}}\gamma_{j_d},$$
$$A_kZ_k' = P^\theta, \qquad\qquad Z_k'/A_k = P^{-(k+1)+k\theta}\gamma_{j_1}\cdot\ldots\cdot\gamma_{j_k}\gamma_{j_d},$$
$$Z_k/Z_k' = P^{1-\theta}/\gamma_{j_k}$$

are satisfied for each $1 \leq k \leq d-1$. Observe that our hypothesis on $\theta$ ensures that $Z_k \geq Z_k'$. Furthermore, one easily confirms that, according to our assumptions on the relative sizes of the $\gamma_{j_k}$, one has

$$Z_k = \sqrt{\frac{\gamma_{j_d}}{\gamma_{j_k}}}\prod_{l=1}^{k-1}\sqrt{\frac{\gamma_{j_l}}{P^{1-\theta}}} \leq 1.$$

It follows that Lemma 3.3 is applicable and yields

$$U_k(Z_k) \ll \left(\frac{Z_k}{Z_k'}\right)^s U_k(Z_k') \ll (P^{1-\theta}/\gamma_{j_k})^s U_k(Z_k'). \tag{3.6}$$

For a given $k$ between 0 and $d-1$ consider the quantity

$$\nu(k) = N_{\mathbf{j}}\left(\underbrace{P^\theta, \ldots, P^\theta}_{\text{first } k \text{ entries}}, \frac{P}{\gamma_{k+1}}, \ldots, \frac{P}{\gamma_{d-1}}; \frac{P^{(k+1)-k\theta}}{\gamma_{j_1}\cdot\ldots\cdot\gamma_{j_k}\gamma_{j_d}}\right), \tag{3.7}$$

then $\nu(k)$ can be expressed in terms of $U_k(Z_k')$ or $U_{k+1}(Z_{k+1})$ by the relation

$$\nu(k) = \sum_{\substack{\mathbf{h}_l \ll P^\theta \\ 1 \leq l \leq k-1}} \sum_{\substack{\mathbf{h}_l \ll P/\gamma_{j_l} \\ k+1 \leq l \leq d-1}} U_k(Z_k') = \sum_{\substack{\mathbf{h}_l \ll P^\theta \\ 1 \leq l \leq k}} \sum_{\substack{\mathbf{h}_l \ll P/\gamma_{j_l} \\ k+2 \leq l \leq d-1}} U_{k+1}(Z_{k+1}). \tag{3.8}$$

On combining (3.6) and (3.8) we obtain a recursive relation for the $\nu(k)$ which is given by $\nu(k-1) \ll (P^{1-\theta}/\gamma_{j_k})^s\nu(k)$. It follows that

$$\nu(d-1) \gg P^{-(1-\theta)(d-1)s}\left(\prod_{k=1}^{d-1}\gamma_{j_k}^s\right)\nu(0),$$

and on recalling the definition (3.7) of $\nu(k)$ we find

$$N_{\mathbf{j}}\left(P^{\theta}, \ldots, P^{\theta}; \frac{P^{d-(d-1)\theta}}{\hat{\gamma}_{\mathbf{j}}}\right) \gg P^{-(d-1)(1-\theta)s}\left(\prod_{k=1}^{d-1} \gamma_{j_k}^s\right) N_{\mathbf{j}}\left(P_{j_1}, \ldots, P_{j_{d-1}}; P_{j_d}\right).$$

Inserting Lemma 3.2 now yields the desired result. $\qquad\square$

The content of Lemma 3.4 is that if the exponential sum is large, the quantities $M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H})$ are simultaneously close to an integer for many choices of $\mathcal{H}$. This is certainly true if the forms $B_n$ tend to vanish for geometric reasons, and in the other case it implies that one can find genuine (i.e. non-zero) solutions to the diophantine approximation problem that is implicit in (3.3). This yields the standard threefold case distinction.

**Lemma 3.5.** *Let $0 < \theta \le 1 - \eta$ and $k > 0$ be parameters, and let $\boldsymbol{\alpha} \in [0,1)^r$. Then one of the following is true.*

*(A) The exponential sum $T(\boldsymbol{\alpha}; \mathcal{P})$ is bounded by*

$$|T(\boldsymbol{\alpha}; \mathcal{P})| \ll \left(\frac{P^m}{\det C}\right)^s P^{-k\theta}.$$

*(B) For every $\mathbf{j} \in J$ one finds integers $(q_{\mathbf{j}}, a_{\mathbf{j}})$ satisfying*

$$0 < q_{\mathbf{j}} \ll P^{(d-1)\theta} \qquad and \qquad |\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \ll P^{-d+(d-1)\theta}\hat{\gamma}_{\mathbf{j}}.$$

*(C) The number of $|\mathbf{h}_l| \le P^{\theta}$ for $1 \le l \le d-1$ that satisfy*

$$B_n(\mathbf{h}_1, \ldots, \mathbf{h}_{d-1}) = 0 \qquad (1 \le n \le s)$$

*is asymptotically greater than $(P^{\theta})^{(d-1)s - 2^{d-1}k - \varepsilon}$.*

*Proof.* The proof is similar to that of [2, Lemma 3.4]. Assuming that the estimate in (A) does not hold, Lemma 3.4 implies that for every $\mathbf{j} \in J$ we have

$$\left\| M(\mathbf{j})\alpha_{\mathbf{j}}B_n(\mathcal{H}_{\mathbf{j}}) \right\| < P^{-d+(d-1)\theta}\hat{\gamma}_{\mathbf{j}} \qquad (1 \le n \le s)$$

for at least $P^{(d-1)s\theta - 2^{d-1}k\theta - \varepsilon}$ choices of $\mathcal{H}_{\mathbf{j}} \le P^{\theta}$. If $B_n(\mathcal{H}_{\mathbf{j}})$ is non-zero for some $n$ and some $\mathcal{H}_{\mathbf{j}}$ counted by $N_{\mathbf{j}}(P^{\theta}, \ldots, P^{\theta}; P^{d-(d-1)\theta}/\hat{\gamma}_{\mathbf{j}})$, we denote its value by $q_{\mathbf{j}}$. Obviously $q_{\mathbf{j}} \ll P^{(d-1)\theta}$, and it follows that we can find an integer $a_{\mathbf{j}}$ with the property that

$$|\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \ll P^{-d+(d-1)\theta}\hat{\gamma}_{\mathbf{j}}.$$

This establishes the statement. $\qquad\square$

By choosing the number of variables large enough, the singular case can be excluded. This is, however, identical to the treatment in [2, Lemma 3.5].

**Lemma 3.6.** *Let $\boldsymbol{\alpha} \in [0,1)^r$ and let $0 < \theta \le 1 - \eta$ and $k$ be parameters with*

$$s - \dim \operatorname{Sing} F > 2^{d-1}k. \tag{3.9}$$

*Then the alternatives are the following.*

*(A) The exponential sum $T(\boldsymbol{\alpha}; \mathcal{P})$ is bounded by*

$$|T(\boldsymbol{\alpha}; \mathcal{P})| \ll \left(\frac{P^m}{\det C}\right)^s P^{-k\theta + \varepsilon}.$$

*(B) For every $\mathbf{j} \in J$ one finds a pair of coprime integers $(q_{\mathbf{j}}, a_{\mathbf{j}})$ satisfying*

$$0 < q_{\mathbf{j}} \ll P^{(d-1)\theta} \qquad and \qquad |\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \ll P^{-d+(d-1)\theta} \hat{\gamma}_{\mathbf{j}}.$$

*Proof.* This is essentially Lemma 3.5 of [2]. Notice that the singular case in Lemma 3.5 is the same as that in Lemma 3.4 of our former work, and in particular does not depend on the matrix $C$ or indeed on linear spaces altogether, so the methods used to derive Lemma 3.5 from Lemma 3.4 of [2] are applicable, and the singular case is excluded by our assumption (3.9). $\qquad\square$

## 4. IMPLEMENTATION OF THE CIRCLE METHOD

Let $c$ be sufficiently large in terms of $C$ and the coefficients of $F$, then we write $\mathfrak{M}(P, \theta)$ for the set of all $\boldsymbol{\alpha} \in [0,1)^r$ that have a rational approximation satisfying

$$0 \leq a_{\mathbf{j}} < q_{\mathbf{j}} \leq cP^{(d-1)\theta} \qquad and \qquad |\alpha_{\mathbf{j}} q_{\mathbf{j}} - a_{\mathbf{j}}| \leq cP^{-d+(d-1)\theta} \hat{\gamma}_{\mathbf{j}}$$

for all $\mathbf{j} \in J$, and $\mathfrak{m}(P, \theta)$ for the complement thereof. Notice that this respects the case distinction of Lemma 3.6, so there is a non-trivial minor arcs estimate for all $\boldsymbol{\alpha} \in \mathfrak{m}(P, \theta)$. In order to keep notation simple, we omit the parameter $P$ whenever there is no danger of confusion.

**Lemma 4.1.** *Suppose the parameters $k$ and $\theta$ satisfy*

$$0 < \theta < \theta_0 = \min\left\{ \frac{d}{2(d-1)}, 1 - \eta \right\}$$

*and*

$$k > \max\left\{ 2r(d-1), rd\frac{\sum_{i=1}^{m} \log P_i}{m \log P_{\min}} \right\}. \tag{4.1}$$

*Then there exists a $\delta > 0$ such that the minor arcs contribution is bounded by*

$$\int_{\mathfrak{m}(\theta)} |T(\boldsymbol{\alpha}; \mathcal{P})| \mathrm{d}\boldsymbol{\alpha} = O\left( \left( \frac{P^m}{\det C} \right)^{s-rd/m} P^{-\delta} \right).$$

*Proof.* We follow the proof of [1, Lemma 4.4]. As a first step, we note that

$$\left( \frac{P^m}{\det C} \right)^{-rd/m} = P^{-rd(1-\Omega)},$$

where $\Omega = \frac{1}{m} \sum_{i=1}^{m} \frac{\log \gamma_i}{\log P}$. It follows that $P^{-k\theta_0} \ll (P^m/\det C)^{-rd/m-\delta}$ as soon as either one has $\theta_0 = \frac{d}{2(d-1)}$ and $k > 2r(d-1)$, or $\theta_0 = 1 - \eta$ and

$$k > rd\frac{1-\Omega}{1-\eta} = rd\frac{m \log P - \sum_{i=1}^{m} \log \gamma_i}{m(\log P - \log \gamma_{\max})} = rd\frac{\sum_{i=1}^{m} \log P_i}{m \log P_{\min}}.$$

It follows that for some $\delta > 0$ the contribution arising from $\mathfrak{m}(\theta_0)$ is given by

$$\int_{\mathfrak{m}(\theta_0)} |T(\boldsymbol{\alpha}; \mathcal{P})| \mathrm{d}\boldsymbol{\alpha} \ll \left( \frac{P^m}{\det C} \right)^{s-rd/m} P^{-\delta}.$$

Given $0 < \theta < \theta_0$, we can find $\delta > 0$ and $1 \geq \theta_0 > \theta_1 > \ldots > \theta_M = \theta > 0$ satisfying

$$(k - 2r(d-1))\theta > 2\delta \quad and \quad (\theta_i - \theta_{i+1})k < \delta \quad for\ all\ i. \tag{4.2}$$

This is always possible with $M = O(1)$. Then on writing

$$\mathfrak{m}_i = \mathfrak{m}(\theta_i) \setminus \mathfrak{m}(\theta_{i-1}) = \mathfrak{M}(\theta_{i-1}) \setminus \mathfrak{M}(\theta_i)$$

a straightforward calculation shows that

$$\operatorname{vol} \mathfrak{m}_i \leq \operatorname{vol} \mathfrak{M}(\theta_{i-1}) \ll \left( \frac{P^m}{\det C} \right)^{-rd/m} P^{2(d-1)r\theta_{i-1}}.$$

Recall that for $\boldsymbol{\alpha} \in \mathfrak{m}(\theta_i)$ we are in the situation of case (A) in Lemma 3.6, so the $i$-th minor arcs contribution is bounded by

$$\int_{\mathfrak{m}_i} |T(\boldsymbol{\alpha}; \mathcal{P})| \, \mathrm{d}\boldsymbol{\alpha} \ll \operatorname{vol} \mathfrak{M}(\theta_{i-1}) \sup_{\boldsymbol{\alpha} \in \mathfrak{m}(\theta_i)} |T(\boldsymbol{\alpha}; \mathcal{P})|$$

$$\ll \left( \frac{P^m}{\det C} \right)^{-rd/m} P^{2(d-1)r\theta_{i-1}} \left( \frac{P^m}{\det C} \right)^{s} P^{-k\theta_i}$$

$$\ll \left( \frac{P^m}{\det C} \right)^{s-rd/m} P^{-k\theta_i + 2r(d-1)\theta_{i-1}}.$$

By (4.2) the exponent is

$$-k\theta_i + 2r(d-1)\theta_{i-1} = k(\theta_{i-1} - \theta_i) - (k - 2r(d-1))\theta_{i-1}$$
$$< k(\theta_{i-1} - \theta_i) - (k - 2r(d-1))\theta < -\delta$$

and on summing the $O(1)$ contributions with $1 \leq i \leq M$ we recover the statement. $\square$

We now have to homogenise the major arcs in order to find a common denominator for the major arcs approximations. However, for our present context it is enough to give only a crude bound here and let $q = \operatorname{lcm}_{\mathbf{j} \in J} q_{\mathbf{j}}$, so that trivially $q \ll P^{r(d-1)\theta}$ and $|\alpha_{\mathbf{j}} q - b_{\mathbf{j}}| \ll P^{-d+r(d-1)\theta} \hat{\gamma}_{\mathbf{j}}$.

For technical reasons, it is convenient to extend the major arcs slightly and define $\mathfrak{M}'(\theta)$ to be the set of all $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$ contained in the interval $[0, 1)^r$ that satisfy

$$0 \leq \mathbf{a} < q \leq c' P^{r(d-1)\theta} \qquad \text{and} \qquad |\beta_{\mathbf{j}}| \leq c' P^{-d+r(d-1)\theta} \hat{\gamma}_{\mathbf{j}} \quad (\mathbf{j} \in J) \qquad (4.3)$$

for some suitably large constant $c'$. Henceforth all parameters $\boldsymbol{\alpha}, \mathbf{a}, q, \boldsymbol{\beta}$ will be implicitly understood to satisfy the major arcs inequalities as given in (4.3). Let

$$S_q(\mathbf{a}) = \sum_{\overline{\mathbf{x}} \pmod q} e\left( \frac{\mathfrak{F}(\overline{\mathbf{x}}; \mathbf{a})}{q} \right)$$

and

$$v_{\mathcal{P}}(\boldsymbol{\beta}) = \int_{\overline{\boldsymbol{\xi}} \in \mathcal{P}^s} e\left( \mathfrak{F}(\overline{\boldsymbol{\xi}}; \boldsymbol{\beta}) \right) \mathrm{d}\overline{\boldsymbol{\xi}}.$$

We can now replace the exponential sum by an expression in terms of the approximation given by $\mathbf{a}$, $q$ and $\boldsymbol{\beta}$.

**Lemma 4.2.** *Assume that* $\boldsymbol{\alpha} = \mathbf{a}/q + \boldsymbol{\beta}$, *then one has*

$$T(\boldsymbol{\alpha}; \mathcal{P}) - q^{-ms} S_q(\mathbf{a}) v_{\mathcal{P}}(\boldsymbol{\beta}) \ll q^{ms} + \left( \frac{P^m}{\det C} \right)^{s} \left( \sum_{\mathbf{j} \in J} |\beta_{\mathbf{j}}| \frac{P^d}{\hat{\gamma}_{\mathbf{j}}} \right) \frac{q}{P_{\min}}.$$

*Proof.* We sort the variables $\overline{\mathbf{x}}$ occurring in $T(\boldsymbol{\alpha}; \mathcal{P})$ in arithmetic progressions modulo $q$ and write $\overline{\mathbf{x}} = q\overline{\mathbf{w}} + \overline{\mathbf{z}}$. Then

$$T(\boldsymbol{\alpha}; \mathcal{P}) - q^{-ms} S_q(\mathbf{a}) v_{\mathcal{P}}(\boldsymbol{\beta}) = \sum_{\overline{\mathbf{z}} \ (\mathrm{mod}\ q)} e(\mathfrak{F}(\overline{\mathbf{z}}; \mathbf{a}/q)) H(q, \boldsymbol{\beta}, \overline{\mathbf{z}}),$$

where

$$H(q, \boldsymbol{\beta}, \overline{\mathbf{z}}) = \sum_{\substack{\overline{\mathbf{w}} \in \mathbb{Z}^{ms} \\ q\overline{\mathbf{w}} + \overline{\mathbf{z}} \in \mathcal{P}^s}} e(\mathfrak{F}(q\overline{\mathbf{w}} + \overline{\mathbf{z}}; \boldsymbol{\beta})) - q^{-ms} \int_{\mathcal{P}^s} e(\mathfrak{F}(\overline{\boldsymbol{\zeta}}; \boldsymbol{\beta})) \mathrm{d}\overline{\boldsymbol{\zeta}}.$$

Including the first term of $H(q, \boldsymbol{\beta}, \overline{\mathbf{z}})$ in the integral, we may apply the Mean Value Theorem and obtain

$$H(q, \boldsymbol{\beta}, \overline{\mathbf{z}}) = \sum_{\substack{\overline{\mathbf{w}} \in \mathbb{Z}^{ms} \\ q\overline{\mathbf{w}} + \overline{\mathbf{z}} \in \mathcal{P}^s}} \int_{\overline{\mathbf{w}}}^{\overline{\mathbf{w}}+1} \left( e(\mathfrak{F}(q\overline{\mathbf{w}} + \overline{\mathbf{z}}; \boldsymbol{\beta})) - e(\mathfrak{F}(q\overline{\boldsymbol{\zeta}} + \overline{\mathbf{z}}; \boldsymbol{\beta})) \right) \mathrm{d}\overline{\boldsymbol{\zeta}} + O(1)$$

$$\ll 1 + \left( \frac{(Pq^{-1})^m}{\det C} \right)^s \left( \sum_{\mathbf{j} \in J} |\beta_{\mathbf{j}}| \frac{P^d}{\hat{\gamma}_{\mathbf{j}}} \right) \frac{q}{P_{\min}}.$$

The remaining term is just $S_q(\mathbf{a})$ and can be bounded trivially by $q^{ms}$, so altogether we have

$$T(\boldsymbol{\alpha}; \mathcal{P}) - q^{-ms} S_q(\mathbf{a}) v_{\mathcal{P}}(\boldsymbol{\beta}) \ll q^{ms} \left( 1 + \left( \frac{(Pq^{-1})^m}{\det C} \right)^s \left( \sum_{\mathbf{j} \in J} |\beta_{\mathbf{j}}| \frac{P^d}{\hat{\gamma}_{\mathbf{j}}} \right) \frac{q}{P_{\min}} \right)$$

as claimed. $\qquad \square$

We define the truncated singular series and singular integral as

$$\mathfrak{S}_{\psi}(P) = \sum_{q=1}^{c' P^{r(d-1)\theta}} q^{-ms} \sum_{\substack{\mathbf{a}=0 \\ (\mathbf{a}, q)=1}}^{q-1} S_q(\mathbf{a}) e\left( \frac{-\mathbf{a} \cdot \mathbf{n}}{q} \right)$$

and

$$\mathfrak{I}_{\psi}(\mathcal{P}) = \int_{|\beta_{\mathbf{j}}| \leq c' P^{-d + r(d-1)\theta} \hat{\gamma}_{\mathbf{j}}} v_{\mathcal{P}}(\boldsymbol{\beta}) e(-\boldsymbol{\beta} \cdot \mathbf{n}) \mathrm{d}\boldsymbol{\beta},$$

respectively. This notation allows us to determine the overall error arising from substituting $q^{-ms} S_q(\mathbf{a}) v(\boldsymbol{\beta})$ for $T(\boldsymbol{\alpha}; \mathcal{P})$ by integrating the expression from Lemma 4.2 over $\mathfrak{M}'(\theta)$.

**Lemma 4.3.** *The total major arcs contribution is given by*

$$\int_{\mathfrak{M}'(P,\theta)} T(\boldsymbol{\alpha}; \mathcal{P}) e(-\boldsymbol{\alpha} \cdot \mathbf{n}) \mathrm{d}\boldsymbol{\alpha} = \mathfrak{S}_{\psi}(P) \mathfrak{I}_{\psi}(\mathcal{P}) + O\left( \left( \frac{P^m}{\det C} \right)^{s - rd/m} P^{(d-1)r(2r+3)\theta - 1 + \eta} \right).$$

*Proof.* We have

$$\mathrm{vol}\, \mathfrak{M}'(\theta) \ll \sum_{q=1}^{c' P^{r(d-1)\theta}} \left( q P^{-d + r(d-1)\theta} \right)^r \prod_{\mathbf{j} \in J} \hat{\gamma}_{\mathbf{j}} \ll \left( \frac{P^m}{\det C} \right)^{-rd/m} P^{(d-1)r(2r+1)\theta},$$

where the last step follows by (2.3). On the other hand, inserting the major arcs inequalities (4.3) for $q$ and $\boldsymbol{\beta}$ in the bound from Lemma 4.2 yields

$$|T(\boldsymbol{\alpha};\mathcal{P}) - q^{-ms}S_q(\mathbf{a})v_{\mathcal{P}}(\boldsymbol{\beta})| \ll \left(\frac{P^m}{\det C}\right)^s P^{-1+2r(d-1)\theta+\eta}.$$

This implies the statement. □

We may now fix $\theta$ such that

$$0 < \theta \le \frac{1 - \eta - \delta}{(d-1)r(2r+3)}$$

holds for some small $\delta > 0$. Under this condition, Lemmata 4.1 and 4.3 can be combined to establish

$$N_\psi(\mathcal{P}) = \mathfrak{S}_\psi(P)\mathfrak{J}_\psi(\mathcal{P}) + O\left(\left(\frac{P^m}{\det C}\right)^{s-rd/m}P^{-\delta}\right).$$

As in [2] the truncated singular series and integral can be extended to infinity. As usual, the expected growth rate is encoded in the geometry of the problem and derives from normalising the singular integral. In fact, standard computations reveal that

$$\mathfrak{J}_\psi(\mathcal{P}) = \left(\frac{P^m}{\det C}\right)^{s-rd/m} \int_{|\boldsymbol{\beta}| \le c'P^{r(d-1)\theta}} v_1(\boldsymbol{\beta})e(-\boldsymbol{\beta}\cdot\tilde{\mathbf{n}})\mathrm{d}\boldsymbol{\beta}, \qquad (4.4)$$

where $\tilde{n}_{\mathbf{j}} = P^{-d}\hat{\gamma}_{\mathbf{j}}n_{\mathbf{j}}$ for every $\mathbf{j} \in J$.

As in [2] the truncated singular series and integral can be extended to infinity. In particular, $v_1(\boldsymbol{\beta})$ and $S_q(\mathbf{a})$ are independent of $C$ and are therefore identical to the respective quantities considered in our former work [2]. We present here a slight refinement of the treatment given in [2], leading to our improved bounds on the number of variables that is required in order to obtain an asymptotic formula. For the sake of reference, we quote here Lemma 3.5 of [2].

**Lemma 4.4.** *Suppose $C = \mathrm{Id}_m$, and $k > 0$ and $0 < \theta \le d/(2(d-1))$ are parameters with $s - \dim \mathrm{Sing}\, F > 2^{d-1}k$. Then one of the following is true.*
*(A) $\boldsymbol{\alpha} \in \mathfrak{m}(Q,\theta)$, i.e. $T(\boldsymbol{\alpha}, Q) \ll Q^{ms-k\theta}$, or*
*(B) $\boldsymbol{\alpha} \in \mathfrak{M}(Q,\theta)$, i.e. for every $\mathbf{j} \in J$ one finds integers $0 \le a_{\mathbf{j}} < q_{\mathbf{j}} \ll Q^{(d-1)\theta}$ and real numbers $|\beta_{\mathbf{j}}| \ll q_{\mathbf{j}}^{-1}Q^{-d+(d-1)\theta}$ satisfying $\alpha_{\mathbf{j}} = a_{\mathbf{j}}/q_{\mathbf{j}} + \beta_{\mathbf{j}}$, and this representation is unique.*

This allows us to understand the contribution from the singular series and the singular integral.

**Lemma 4.5.** *The function $S_q(\mathbf{a})$ is pseudo-multiplicative, i.e.*

$$S_q(\mathbf{a})S_{q'}(\mathbf{a}') = S_{qq'}(q'\mathbf{a} + q\mathbf{a}') \qquad (4.5)$$

*whenever $(q, q') = 1$. Furthermore, when $k > (d-1)(r+1+\delta)$ one has*

$$q^{-ms}S_q(\mathbf{a}) \ll q^{-(r+1+\delta)} \qquad (4.6)$$

*for every $\mathbf{a} \pmod{q}$ with $(q, \mathbf{a}) = 1$, and*

$$v_1(\boldsymbol{\gamma}) \ll |\boldsymbol{\gamma}|^{-(r+1+\delta)}. \qquad (4.7)$$

*Proof.* The multiplicativity follows by standard arguments, and the second statement is [2, Lemma 6.3] with $W = r + 1 + \delta$. Observe that in [2], only $d \geq 3$ is permitted, but a slight modification of the argument shows that the statement remains valid for $d = 2$ as well.

In the case of $v_1(\boldsymbol{\gamma})$ the argument is similar. Let $\theta$ be given as above and choose $Q$ in such a way that $|\boldsymbol{\gamma}| = cQ^{(d-1)\theta}$. As before, the major arcs are disjoint, and the pont $P^{-d}\boldsymbol{\gamma}$ lies right on the edge with the approximation $\mathbf{a} = \mathbf{0}$ and $q = 1$. It therefore follows from Lemma 4.3 that

$$v_1(\boldsymbol{\gamma}) \ll Q^{-ms}T(Q^{-d}\boldsymbol{\gamma}, Q) + Q^{-1}|\boldsymbol{\gamma}|.$$

As before, by continuity the the minor arcs bound continues to apply on the boundaries of $\mathfrak{m}(Q, \theta)$ and yields a bound $T(Q^{-d}\boldsymbol{\gamma}, Q) \ll Q^{ms-k\theta}$. Altogether we thus obtain

$$v_1(\boldsymbol{\gamma}) \ll Q^{-k\theta} + Q^{-1+(d-1)\theta} \ll |\boldsymbol{\gamma}|^{-(r+1+\delta)}.$$

This completes the proof of the lemma. $\qquad\square$

Now define the complete singular series $\mathfrak{S}_\psi$ and the singular integral $\mathfrak{J}_\psi$ as

$$\mathfrak{S}_\psi = \sum_{q=1}^{\infty} \sum_{\substack{\mathbf{a}=0 \\ (\mathbf{a},q)=1}}^{q-1} q^{-ms}S_q(\mathbf{a})e\left(\frac{-\mathbf{a}\cdot\mathbf{n}}{q}\right)$$

and

$$\mathfrak{J}_\psi = \int_{\mathbb{R}^r} v_{\mathcal{P}}(\boldsymbol{\beta})e(-\boldsymbol{\beta}\cdot\mathbf{n})\mathrm{d}\boldsymbol{\beta}.$$

Then we have shown that

$$N_\psi(\mathcal{P}) = \mathfrak{J}_\psi\mathfrak{S}_\psi + O\left(\left(\frac{P^m}{\det C}\right)^{s-rd/m}P^{-\delta}\right),$$

where the bound on the number of variables is obtained from combining (3.9) and (4.1) as well as the bound on $k$ given in Lemma 4.5 and is given by

$$s - \dim \mathrm{Sing}\, F > 2^{d-1}\max\left\{2(d-1), rd\frac{\sum_{i=1}^{m}\log P_i}{m\log P_{\min}}\right\},$$

which is the bound claimed in Theorem 1.2.

It follows from (4.5) that the singular series can be expanded as an Euler product

$$\mathfrak{S}_\psi = \prod_p \chi_p(F; \psi)$$

whose factors are given by

$$\chi_p(F; \psi) = \sum_{l=0}^{\infty} p^{-lms} \sum_{\substack{\mathbf{a}=0 \\ (\mathbf{a},p)=1}}^{p^l-1} S_{p^l}(\mathbf{a})e\left(\frac{-\mathbf{a}\cdot\mathbf{n}}{p^l}\right).$$

In particular, (4.6) implies that $\chi_p(F; \psi) = 1 + O(p^{-1-\delta})$, whence we may conclude that the Euler product converges and thus vanishes if and only if one of the factors

is zero. Standard arguments now show that

$$\chi_p(F;\psi) = \lim_{l\to\infty}(p^l)^{r-ms}\,\mathrm{Card}\{\overline{\mathbf{x}} \pmod{p^l} : \Phi_{\mathbf{j}}(\overline{\mathbf{x}}) \equiv n_{\mathbf{j}} \pmod{p^l} \quad \forall\,\mathbf{j}\in J\}$$

$$= \lim_{l\to\infty}(p^l)^{r-ms}\,\mathrm{Card}\{\overline{\mathbf{x}} \pmod{p^l} :$$

$$F(\mathbf{x}_1 t_1 + \cdots + \mathbf{x}_m t_m) \equiv \psi(t_1,\ldots,t_m) \pmod{p^l}\}$$

identically in $t_1,\ldots,t_m$, where we used the correspondence between identical representations of forms and points on hypersurfaces as in §2. Thus $\chi_p(F;\psi)$ measures the density of solutions to (1.2) in $\mathbb{Q}_p$, and it follows from [3, Theorem 5.1] in combination with [8, Theorem 4.1] that this density is positive for all $p$ whenever

$$ms - \dim\mathrm{Sing}_m F \geq (d^2+m)^{2^{d-2}}d^{2^{d-1}},$$

where

$$\mathrm{Sing}_m F = \left\{\overline{\mathbf{x}} : \mathrm{rank}\left(\frac{\partial\Phi_{\mathbf{j}}(\overline{\mathbf{x}})}{\partial\mathbf{x}_i}\right)_{\mathbf{j}\in J, 1\leq i\leq m} \leq r-1\right\},$$

and one probably has $\dim\mathrm{Sing}_m F = (m-1)s + \dim\mathrm{Sing}\, F$.

Similarly, standard computations show that the singular integral is given by

$$\mathfrak{J}_\psi = \left(\frac{P^m}{\det C}\right)^{s-rd/m}\chi_\infty(\psi;P_1,\ldots,P_m),$$

where the integral

$$\chi_\infty(\psi;P_1,\ldots,P_m) = \int_{\mathbb{R}^r}\int_{|\overline{\boldsymbol{\xi}}|\leq 1} e\left(\mathfrak{F}(\overline{\boldsymbol{\xi}};\boldsymbol{\beta}) - \boldsymbol{\beta}\cdot\tilde{\mathbf{n}}\right)\mathrm{d}\overline{\boldsymbol{\xi}}\mathrm{d}\boldsymbol{\beta} \tag{4.8}$$

measures solutions of a rescaled version of the problem in the real unit box. In fact, a straightforward generalisation of §11 and Lemma 2 of [7] shows that

$$\chi_\infty(\psi;P_1,\ldots,P_m) \gg 1$$

whenever $s - \dim\mathrm{Sing}\, F > 2^d r(d-1)$ and the variety

$$M = \{\overline{\mathbf{x}}\in(-1,1)^{ms} : \Phi_{\mathbf{j}}(\overline{\mathbf{x}}) = \tilde{n}_{\mathbf{j}} \quad (\mathbf{j}\in J)\} \subset \mathbb{R}^{ms} \tag{4.9}$$

is of dimension at least $ms - r$. As above, this variety may be rewritten in terms of $F$ and $\psi$ themselves. In fact, one has

$$\sum_{\mathbf{j}\in J}\tilde{n}_{\mathbf{j}}t_{j_1}\cdot\ldots\cdot t_{j_d} = \sum_{\mathbf{j}\in J}P^{-d}n_{\mathbf{j}}\gamma_{j_1}t_{j_1}\cdot\ldots\cdot\gamma_{j_d}t_{j_d}$$

$$= \sum_{\mathbf{j}\in J}(t_{j_1}/P_{j_1})\cdot\ldots\cdot(t_{j_d}/P_{j_d}) = \psi(\tfrac{t_1}{P_1},\ldots,\tfrac{t_m}{P_m}),$$

whence $M$ is given by

$$M = \{\overline{\mathbf{x}}\in(-1,1)^{ms} : F(\mathbf{x}_1 t_1 + \cdots + \mathbf{x}_m t_m) = \psi(\tfrac{t_1}{P_1},\ldots,\tfrac{t_m}{P_m}) \quad \forall\, t_1,\ldots,t_m\},$$

and the requirement is that this variety have a positive $(ms-r)$-dimensional volume. This condition will be satisfied as soon as $M$ contains at least one non-singular point.

Whether or not $M$ contains any points depends on the choice of the parameters $P_i$ as well as the shape of $\psi$ itself. In the case of Theorem 1.1 we pick $P_i = n_i^{1/d}$, so

$$\psi(\tfrac{t_1}{P_1},\ldots,\tfrac{t_m}{P_m}) = \widetilde{\psi}(t_1,\ldots,t_m)$$

where $\widetilde{\psi}$ is as in (2.1) With this choice, we have $\tilde{n}_i = 1$ for $1 \le i \le m$, and $\tilde{n}_{\mathbf{j}} \le 1$ for all $\mathbf{j} \in J$ if $\psi$ is pseudo-diagonal. Observe that $F$ maps the unit cube to an area of size $O(1)$, so $M$ will not contain any points unless all coefficients of $\tilde{\psi}$ are at most of size 1. In other words, if $\psi$ is not pseudo-diagonal up to a factor at most as large as the largest coefficient of $F$, the variety $M$ will invariably be empty.

## References

[1] B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/1962), 245–263.

[2] J. Brandes, *Forms representing forms and linear spaces on hypersurfaces*, Proc. London Math. Soc. **108** (2014), 809–835.

[3] J. Brandes, *A note on p-adic solubility for forms in many variables*, Bull. London Math. Soc. **47** (2015), 501–508.

[4] H. Davenport, Analytic Methods for Diophantine Equations and Diophantine Inequalities, Cambridge University Press, Cambridge, 2005.

[5] R. Dietmann and M. Harvey, *On the representation of quadratic forms by quadratic forms*, Michigan Math. J. **62** (2013), 869–889.

[6] S. Raghavan, *Modular forms of degree n and representation by quadratic forms*, Ann. of Math. (2) **70** (1959), 446–477.

[7] W. M. Schmidt, emphSimultaneous rational zeros of quadratic forms, Seminar on Number Theory, Paris 1980-81 (Paris, 1980/1981). In: Progress in Mathematics **22**, Birkhäuser, Boston, MA, 1982, pp. 281–307.

[8] T. D. Wooley, *On the local solubility of diophantine systems*, Compositio Math. **111** (1998), 149–165.

Mathematisches Institut, Bunsenstr. 3–5, 37073 Göttingen, Germany
*E-mail address*: jbrande@uni-math.gwdg.de